Natural Language Generation from Pictographs

Leen Sevens, Vincent Vandeghinste, Ineke Schuurman, Frank Van Eynde
Centrum voor Computerlinguïstiek KU Leuven
firstname@ccl.kuleuven.be

Being unable to access ICT is a major form of social exclusion. For people with Intellectual or Developmental Disabilities (IDD), the use of social media or applications that require the user to be able to read or write well, such as email clients, is a huge stumbling block if no personal assistance is given. There is a need for digital communication interfaces that enable written contact for people with IDD.

We present a first version of a Pictograph-to-Text translation system. It provides help in constructing Dutch textual messages by allowing the user to introduce a series of pictographs and translates these messages into natural language using WordNet synsets and a trigram language model. English and Spanish versions of the tool are currently in development. It can be considered as the inverse translation engine of the Text-to-Pictograph system as described by Vandeghinste et al. (in press), which is primarily conceived to improve the comprehension of textual content.

The Pictograph-to-Text translation engine relies on pictograph input. We have developed two different input methods. The first approach offers a static hierarchy of pictographs, while the second option scans the user input and dynamically adapts itself in order to suggest appropriate pictographs. Two different prototypes for this pictograph predictor have been developed so far.

The first evaluations show that a trigram language model for finding the most likely combination of every pictograph's alternative textual representations is already an improvement over the initial baseline, but there is ample room for improvement in future work.

Reference:
Vandeghinste, Schuurman, Sevens & Van Eynde (in press). Translating Text into Pictographs. Natural Language Engineering. Cambridge University Press.